RS
5
8-15-02

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

ATTY.'S DOCKET: REINSCHMIDT=1

In re Application of:	)	Confirmation No. 7677
Mendel M. REINSHMIDT et al	)	Art Unit: 2661
Appln. No.: 10/091,590	)	Examiner:
Filed: March 7, 2002	)	Washington, D.C.
For: METHOD AND SYSTEM FOR PROVIDING AN IMPROVED...	)	July 17, 2002

REQUEST FOR PRIORITY

Honorable Commissioner for Patents
Washington, D.C. 20231

Sir:

In accordance with the provisions of 37 CFR §1.55 and the requirements of 35 U.S.C. §119, filed herewith a certified copy of:

Israel Appln. No.:141855 Filed: March 7, 2001 .

It is respectfully requested that applicant be granted the benefit of the priority date of the foreign application.

Respectfully submitted,

BROWDY AND NEIMARK, P.L.L.C.
Attorneys for Applicant(s)

By _____
Norman J. Latker
Registration No. 19,963

NJL:jmb
Telephone No.: (202) 628-5197
Facsimile No.: (202) 737-3528
G:/bn/l/luzz/reinshmidt1/priorityPTO.doc

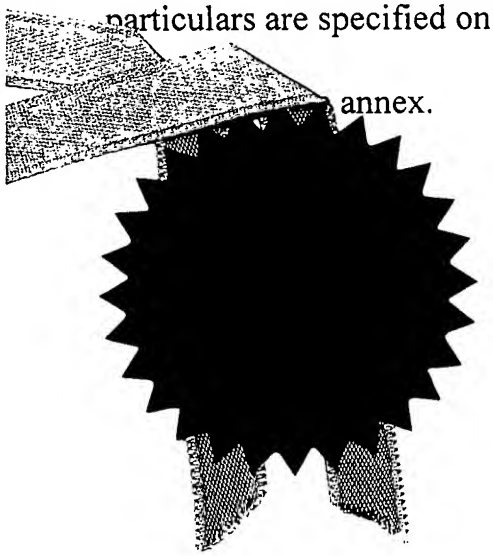
מדינת ישראל
STATE OF ISRAEL

Ministry of Justice
Patent Office

משרד המשפטים
לשכת הפטנטים

This is to certify that annexed hereto is a true copy of the documents as originally deposited with the patent application of which particulars are specified on the annex.

זאת לתעודה כי רצופים בזה העתקים נכונים של המסמכים שהופקדו לכתחילה עם הבקשה לפטנט לפי הפרטים הרשומים בעמוד הראשון של הנספח.



This 23-04-2002 היום

רשם הפטנטים
Commissioner of Patents

נתאשר
Certified

CERTIFIED COPY OF PRIORITY DOCUMENT

חוק הפטנטים, תשכ"ז-1967
Patent Law, 5727 - 1967

בקשה לפטנט
Application for Patent

לשימוש הלישכה
For Office Use

מספר: Number	141855
תאריך: Date	07 -03- 2001
הוקדם/נדחה Ante/Post-Dated	

אני, (שם המבקש, מענו ולגבי גוף מאוגד - מקום התאגדותו)
I, (Name and address of applicant, and in case of body corporate-place of incorporation)

OneTierCommunications, Inc.
Centerville Road
Suite 400
Wilmington, Delaware
County of New Castle 19808
United States of America

Inventor: הממציא:
1. מנחם רינשמיט

Menachem REINSHMIDT

בעל ההמצאה מכח THE LAW הדין ששמה הוא
Owner, by virtue of an invention the title of which is

(בעברית) (Hebrew)
שיטה ומכשיר לקבלת איכות משופרת של שירות להעברת מידע באינטרנט

(באנגלית) (English)
A METHOD AND APPARATUS FOR PROVIDING AN IMPROVED QUALITY OF SERVICE FOR DATA TRANSFER OVER THE INTERNET

hereby apply for a patent to be granted to me in respect thereof. מבקש בזאת כי ינתן לי עליה פטנט

*בקשת חלוקה - Application of Division	*בקשת פטנט מוסף - Application for Patent Addition	*דרישה דין קדימה Priority Claim		
*מבקשת פטנט from Application	*לבקשה/לפטנט to Patent/Appl.	מספר/סימן Number/Mark	תאריך Date	מדינת האיגוד Convention Country
מס' Nº _____ dated מיום _____	מס' Nº _____ dated מיום _____			

*יפוי כח: כללי / ~~מיוחד - רצוף בזה~~ /עוד יוגש
P.O.A.: general /~~individual - attached~~ / to be filed later
filed in case הוגש בענין _____

המען למסירת מסמכים בישראל
Address for Service in Israel
לוצאטו את לוצאטו
ת.ד. 5352
באר שבע 84152
מספרנו: 12145/00

היום 7 בחודש במרץ שנה 2001
This of the year

חתימת המבקש
Signature of Applicant

Luzzatto & Luzzatto

By:
Attorneys for Applicant

לשימוש הלישכה

משרד הפטנטים Patent Office
Ministry Of Justice משרד המשפטים

טופס זה כשהוא מוטבע בחותם לישכת הפטנטים ומושלם במספר ובתאריך ההגשה, הינו אישור להגשת הבקשה שפרטיה רשומים לעיל.
This form, impressed with the Seal of the Patent Office and indicating the number and date of filing, certifies the filing of the application the particulars of which are set out above.

*מחק את המיותר Delete whatever is inapplicable

Ref: 12145/00

שיטה ומכשיר לקבלת איכות משופרת של שירות להעברת מידע באינטרנט

A METHOD AND APPARATUS FOR PROVIDING AN IMPROVED QUALITY OF SERVICE FOR DATA TRANSFER OVER THE INTERNET

A METHOD AND APPARATUS FOR PROVIDING AN IMPROVED QUALITY OF SERVICE FOR DATA TRANSFER OVER THE INTERNET

Field of the Invention

The present invention relates to the field of data communication. More particularly, the invention relates to a method and apparatus for improving the Quality of Service (QoS) for the transportation of selected data packets over the Internet infrastructure, or other multi autonomous systems networks.

Background of the Invention

The Internet system allows users access to different sources of data and also to send and receive various types of data to/from other users. Theoretically, the best possible network is a network, wherein there is only one global manager, and data routes are dynamically changing to accommodate to congestion status. Such a theoretical network would dynamically divert data packets as soon as a congestion state is about to occur. Because of the way the Internet today is extended, developed and managed, rather than having the desired network behavior, it has the following drawbacks:

1. Routers congestion – this problem is reflected in data packets being delayed or lost, and it becomes acute in times of "rush hours", when numerous Internet users start to surf at the same time. This problem arises, because the Internet is a quasi-static network, namely the Internet network has, in general, a limited dynamic behavior. Congested routers are not bypassed, as data routes are changed only

when a total collapse occurs in the network. Routers are capable of working according to several modes of operation. By choosing, theoretically, the most appropriate mode, it has the potential to dynamically re-route information within an AS. Additionally, links within an AS may change to allow using the best possible route for each data packet. Nevertheless, routers are not fully exploited, since they are configured to work with limited dynamic behavior, because full dynamic behavior causes to instability in the network in terms of routing decisions.

2. The Internet comprises many different Autonomous Systems (ASs), each one has a different routing policy and protocol. Each individual AS is still managed rather efficiently by its owner Internet Service Provider (ISP), in comparison to the Internet network as a whole, because the ISP controls, to some degree, the routes in which data is forwarded in his AS(s). The said relative efficiency of each AS is accomplished by using an Interior Gateway Protocol (IGP). Most of the ASs are implemented by using different kinds of protocols, and by using routers, which have been configured to operate in different modes. In most cases, data packets are required to be forwarded through several ASs that are owned by different ISPs. Therefore, the borderline between each two adjacent ASs is the weakest link in the Internet network, in terms of routing. No matter how efficient the data transportation is in each AS, the problem lays in the borderline between each two ASs, namely neighboring ASs do not efficiently cooperate with each other. Each time data packets have to exit one AS and enter another AS they are handled inefficiently. In order to alleviate the problem of forwarding data from one AS to another AS, an Exterior Gateway Protocol (EGP) is used. More recent version of this protocol is the Border Gateway Protocol (BGP-4, BGP version 4).

Although this protocol was specially designed to 'smooth' transportation of dat a from one AS to another AS by using their dynamic capabilities, it was found that the dynamic behavior of the said protocols contributes to vibrations and instability in the routing mechanism. Therefore, these protocols are reduced to be quasi-static, with all the accompanying failings.

3. Each ISP has a different Data Exchange Policy (DEP). DEP refers to the commercial aspect of cooperation between two, or more, ISPs. According to such a DEP, an ISP may, or may not, use other ISP's routers to forward its own data packets more efficiently. In many cases, an ISP will not be able to use the best routers, even if they are available/free, simply because he has not signed up an agreement with the ISP who owns the specific free routers. As a result of this, such an ISP will have to send data packets by using longer and/or slower routes. In other words, ISPs agreements sometimes impose non-optimal routes, simply because of economical considerations.

4. Internet Network Management – as can be appreciated from the above paragraphs, the Internet is not a manageable network, since different and independent ISPs own and manage different parts of the network. Therefore, the Internet system can not offer an end-to-end policy or end-to-end control or optimization.

One of the most important parameters related to the transportation of data packets over a data network, is the Quality of Service (QoS). The Internet has poor QoS, since it has 'unreliable ' or unexpected nature. Consequently, Internet users are limited to applications that do not require high level of QoS. Additionally, since the form and rate of data flow over the Internet are not controllable or predictable, Internet Service

Providers (ISPs) can not provide their users with a sufficient QoS for specific applications (such as voice and multimedia applications), and therefore the services that can be provided are limited.

Special attention is drawn today to the need to use Internet infrastructure to transfer audio/voice signals, and to enable live conversations, like in a PSTN (Public Switching Telephone Network), and multimedia applications. Data packets, except for voice and video packets, are not sensitive to delay in a sense that even when such a data is delayed, the original data integrity is maintained. On the other hand, Voice and video applications are very sensitive to delay in general, and to delay changes (i.e. jitter) in particular, since synchronization between data transmission and data reception is required. If the level of jitter is high, the original information is highly distorted and can not be successfully reconstructed. As a consequence, Internet users may randomly suffer from poor quality of communication, resulting mainly in substantial dynamically changing delay of packets, low data rate and even packet losses. Access time is prolonged, and communication channels become slower. Under severe conditions, communication is even aborted, and a second attempt must be made to restore communication.

The solution of the above-mentioned problems is partly obtained by protocols such as Resource Reservation Protocol (RSVP) and MultiProtocol Label Switching (MPLS). Such solutions provide the subscriber a good, steady and satisfying QoS. However, the problem with this kind of solution is, that it offers no end-to-end management or control, and it is extremely expensive, and therefore, it is not broadly used.

All of the methods described above have not yet provided satisfactory solutions to the problem of incapability to provide an improved QoS, when

it comes to IP applications with broad commercial usage, particularly voice and multimedia applications. Moreover, the problem of poor and unreliable QoS becomes crucial, as it becomes a severe bottleneck when trying to implement the enormous potential of Internet (i.e. telephony, video, multimedia, data, VPN, e-commerce, internet etc.).

It is an object of the present invention to provide a method for providing an improved QoS for the transportation of selected data packets over the Internet network.

It is another object of the present invention to provide a method and apparatus for improving data transportation from one autonomous system to other autonomous systems, in multiple-autonomous systems, which have no end-to-end routing policy.

It is another object of the present invention to provide a method and apparatus for providing an improved QoS for the transportation of selected data packets over a data network that reduces jitter, caused by the network's infrastructure, in order to allow operating jitter-sensitive applications, such as IP Telephony, video and multimedia communications, using the existing Internet infrastructure

It is another object of the present invention to provide a method and apparatus for providing an improved QoS for the transportation of selected data packets over a data network with reduced delay.

It is another object of the present invention to provide a method and apparatus that allow reducing the cost of telephonic services.

It is still another object of the present invention to provide an improved transfer rate of data packets over an existing infrastructure of IP networks.

Other objects and advantages of the invention will become apparent as the description proceeds.

Summary of the Invention

The invention is directed to a method for improving the quality of transportation of selected data packets over a data network, such as the Internet. Selected nodes which are access points to the data network determining, such that each node may be a source, from which the selected data packets can be sent, or a destination to which the selected data packets can be intended. A plurality of intermediate nodes between the source and the destination is selected, for generating a plurality of alternative paths that consist of segments that are used for routing the selected data packets. The packet transportation parameters are periodically sampled in the segments of each preselected path, each time by sending a plurality of test packets from the source to the destination, along the preselected different paths that are defined by different intermediate nodes and their corresponding interconnecting segments. One or more optimal paths for delivering the selected data packets from the source to the destination are defined according to the transportation parameters and optionally, also according to predefined parameters characterizing the segments by selecting a combination of segments, connected to nodes, with the optimal sampled transportation parameters and/or predefined parameters, that connects the source to the destination. A modified header that contains a sequence of consecutive addresses that correspond to consecutive nodes along an optimal path is generated for each selected data packet and attaching the modified header to the

selected data packet. Each selected data packet is forwarded from the source to the destination along the optimal path, while at each node along the optimal path, starting from the source, the following steps are performed: The modified header is processed and the address that corresponds to the next consecutive node is extracted. The selected data packets are forwarded from the node to its consecutive node using the extracted address. This process is repeated for all nodes until the destination node. At the destination node, the modified header is removing from the selected data packet and whenever desired, its original header is consequently used. One or more nodes may be used as intermediate nodes.

The transportation parameters may include the delay time of data packets from source to destination; the variation of the delay time; and loss of packets. Data may be concurrently delivered from a source to a destination over several paths, and optionally by using weighted distribution of data between paths. The weighted distribution can be determined according to the desired level of QoS between the source and the destination.

The definition of the optimal path may be carried out by measuring and storing the time and/or the order of arrival of test packets through different paths from the source to the destination. The definition of each optimal path from the source to the destination may be dynamically varied, according to the sampling results. Whenever a new optimal path is defined, data packets are sent from the source to the destination over the new optimal path. The optimal path may consist of direct connection between the source and the destination. The predefined parameters may include cost; availability; or agreements with ISPs. A grade may be assigned to each optimal path according to the sampled transportation parameters and/or predefined parameters that correspond to a required QoS and/or to the type of data packets to be sent from the source to the

destination. The grade of at least one optimal path may be dynamically varied according to the sampled transportation parameters and/or to the type of data packets to be sent from the source to the destination.. The transportation of data and/or voice packets from the source to the destination may be split between two or more optimal paths.

The invention is also directed to a data network having improved quality of transportation of selected data packets that comprises:

a) a plurality of nodes being access points to the data network, each of which may be a source from which the selected data packets can be sent, or a destination to which the selected data packets can be intended;

b) a plurality of intermediate nodes between the source and the destination, for generating a plurality of alternative paths, consisting of segments, for routing the selected data packets;

c) at one or more nodes and/or intermediate nodes, circuitry for sending a plurality of test packets from the source to the destination, along the preselected different paths defined by different intermediate nodes and their corresponding interconnecting segments;

d) processing means for defining one or more optimal paths for delivering the selected data packets from the source to the destination according to the transportation parameters and optionally, also according to predefined parameters characterizing the segments and/or nodes, and for selecting a combination of segments, connected to nodes, and having the optimal sampled transportation parameters and/or predefined parameters, that connects the source to the destination;

e) at each source, processing means for generating a modified header, for each selected data packet, that contains a sequence of

consecutive addresses that correspond to consecutive nodes along an optimal path and attaching the modified header to the selected data packet;

f) at each node along the optimal path, starting from the source:

f.1) processing means for processing the modified header and for extracting the address that corresponds to the next consecutive node;

f.2)circuitry for forwarding the selected data packet from the node to its consecutive node along the optimal path using the extracted address; and

g) at the destination node, processing means for removing the modified header from the selected data packet and for obtaining the original header of the selected data packet.

Further processing means may be included for dynamically varying the definition of each optimal path from the source to the destination, according to the sampling results and for sending data packets from the source to the destination over the new optimal path.

If a grade is assigned to each optimal path, the data network may further include processing means for dynamically varying the grade of at least one optimal path according to the sampled transportation parameters and/or to the type of data packets to be sent from the source to the destination. Further processing means may be employed for splitting the transportation of data packets from the source to the destination between two or more optimal paths, such that more transportation is directed to, and distributed between, optimal paths having higher grades than the remaining optimal paths, and less transportation is directed to, and distributed between, the remaining optimal paths.

Brief Description of the Drawings

The above and other characteristics and advantages of the invention will be better understood through the following illustrative and non-limitative detailed description of preferred embodiments thereof, with reference to the appended drawings, wherein:

- Fig. 1 schematically illustrates current Internet routing in the existing Internet network (prior art);
- Fig. 2 schematically illustrates Internet routing, according to a preferred embodiment of the invention;
- Fig. 3 schematically illustrates an originator packet routing, according to a preferred embodiment of the invention;
- Fig. 4 schematically illustrates a intermediate packet routing, according to a preferred embodiment of the invention;
- Fig. 5 schematically illustrates a terminating packet routing, according to a preferred embodiment of the invention;
- Fig. 6A schematically shows the functional overview of the concept, according to a preferred embodiment of the invention;
- Fig. 6B schematically illustrates an example for possible QPing paths from originating node (A) to terminating node (E), according to a preferred embodiment of the invention;
- Fig. 6C schematically illustrates the reversed QResponse path, in response to the Qping packets, according to a preferred embodiment of the invention;
- Fig. 6D schematically illustrates a data packet as it is originated by originating switch, passes through an intermediate switch, and finally terminated at the terminating switch, according to a preferred embodiment of the invention;
- Fig. 7 schematically illustrates an example for possible QPing paths from originating node (A) to terminating node (E), according to a preferred embodiment of the invention;

Fig. 8 schematically illustrates an example for optimizing paths between nodes, according to a preferred embodiment of the invention;

- Fig. 9 schematically illustrates a Qping packet structure, according to a preferred embodiment of the invention;
- Fig. 10 schematically illustrates a Qresponse packet structure, according to a preferred embodiment of the invention; and
- Fig. 11 schematically illustrates a Qdata packet structure, according to a preferred embodiment of the invention.

Detailed Description of Preferred Embodiments

Fig. 1 illustrates a conventional packet routing in the Internet. As can be seen in the figure, two end users (10, 11) may communicate with each other by connecting themselves to the Internet (12) via an ISP (13). In such a network the routing is static or in the best case quasi-static. This means that all data of the end users (and probably other end users connected to same ISPs) shall be routed via the same path (14), resulting in congestion in the routers located along a specific (highlighted) path. At the same time, other routers in Fig. 1 (e.g. 15, 16 and 17) may be unloaded. However, the routing scheme can not use them due to its quasi-static attributes. Therefore, other alternative paths, between the two end users, having much better performance are not exploited.

According to a preferred embodiment of the invention, the problems illustrated in Fig. 1 are solved, according to the scheme illustrated in Fig. 2. Several dedicated Qnodes (a Qnode is a "node" being a selected access point to the Internet, characterized by a pair of exit from/entrance into the Internet, such that arriving data packets can be processed at the node and forwarded to another node and/or to an ISP) are deployed over the Internet, preferably in pre-selected "strategic" points. A strategic point is,

according to the invention, a point that interfaces between an ISP and the Internet, and/or a point at a location that allows an efficient transfer of data packages. Particularly, a strategic point may also be determined according to other considerations, such as commercial and/or technical considerations.

As can be seen in Fig. 2, there are two types of Qnodes. The first one is referred to as an access node (e.g. 21), and it interfaces between an ISP server (13) and the Internet, or between a user and the Internet. Access node can be a source node (e.g. 21), whenever it delivers selected packets from an end-user into the Internet, or a destination node, whenever it delivers selected packets from the Internet to an end-user. The second type of a Qnode is referred to as an intermediate node, and it is used as a backbone (e.g. 23 to 26) of the sub-network that comprises all of the Qnodes and their interconnections (hereinafter called "QVPN"). It should be noted that each access node might also be used as an intermediate node.

Fig. 2 illustrates two possible paths between the two end users (10, 11), one is marked with black bold arrows (part of the path is numbered 27), and a second path marked with thin arrows (partly numbered 28). These paths are an example of optimal paths, according to the invention. As can be seen, both paths differ from the path that is normally selected by conventional (quasi-static) routing scheme as shown in Fig. 1. As can be seen, there are, along the path (Fig. 2), two access Qnodes (21, 22), and two intermediate Qnodes (e.g. 23, 24 in path 27), which create a 'three hops' path (e.g. 21-23, 23-24, 24-22). A 'hop' is defined as a segment (part of the path), which connects two Qnodes along a selected path.

Packets are sent from a source, along an optimal path (e.g. 27), to the final destination (10). This is carried out by using a software, that is executed by a dedicated server in each Qnode, which determines the optimal path by selecting a set of segments connecting between Qnodes over the Internet infrastructure. One segment, for example, connects node 23 to node 24 (Fig. 2).

Fig. 2 depicts only two possible paths. However, according to the invention, other paths may be created. For example, a path may comprise Qnodes 21, 25, 26, 24 and 22. The more there are Qnodes, the more paths there are to select from for the QVPN.

Fig. 2 also depicts end users (10, 11) that are connected to Qnodes (21, 22) through their ISPs (13). However, Qnodes may be deployed in two other ways:

1. Private corporations may install at their site a dedicated Qnode to provide routing capability to a portion or all of their Internet and VPN requirements. This type of Qnode is never used as an intermediate switch/node, but only as a user connection point.
2. Intermediate nodes may be deployed at strategic locations/points on the Internet backbone for the purpose of enhancing the transmission over the network. In this case, these nodes will only be used as intermediate nodes (as shown in Fig. 2, reference numeral 24).

ISP servers guarantee that the originating nodes will accept only QVPN subscribers' data packets. This discrimination is carried out in the ISP router, by routing subscribers' packets to the nearest Qnode, while others are directed to the Internet network in the ordinary way, thus bypassing the Qnode.

Fig. 3 illustrates what happens to an IP packet that reaches a Qnode (31). The Camod (the kernel driver software) driver identifies whether this node is the originator node, by verifying if the IP destination address is different from the Qnode IP address. Since the integrity of the original packet must be kept while traveling along the QVPN path, the original packet is forwarded directly from the driver stage to the Qflow application (i.e. bypassing the IP and TCP levels). The CAMON is the interface between the driver and the QFLOW application. Since the originator Qnode contains the preferred/selected optimized route to the final destination (i.e. the destination Qnode), its corresponding Qflow application level adds a new header (32) to the original packet (33), which contains the relevant selected optimized path details, including intermediate nodes - if there are any. Referring also to Fig. 11, the Qflow level implants an offset number into the header (i.e. hopping number), so that the next nodes, along the selected path, will be able to recognize whether the packet is to be forwarded to the next intermediate end station, or it has arrived to the terminating node. This decision is carried out by comparing the said offset number to the current hop number, which is updated every time the packet enters a node. If the offset number and the current hop number differ, the node puts the next consecutive (intermediate) node's IP address, that the packet should be forwarded to, as the next intermediate end station, in front of the packet, and updates the current number. The modified packet is then transmitted via the local ISP default router (34), into the Internet (35). The Internet routers react to the next Qnode address that was implanted into the packet , as if it was the final destination. The User Datagram Protocol (UDP) is used to achieve high-speed end-to-end transmission. It should be emphasized, that the Internet routers always handle (i.e. route) the packets according to next Qnode destination, namely the routers do not have access to the real final destination.

Fig. 4 illustrates what happens to an IP packet which enters an intermediate Qnode (i.e. relay node 41). The CAMOD identified the packet as explained hereinbefore. The packet is forwarded into the IP level, where the current node IP address (42) is extracted. Referring also to Fig. 11, the Qflow level compares the offset number to the current hop number, and since they don't match, this indicates that the said node is only an intermediate node. Therefore, the Qflow level updates the current hop number and inserts in the packet's header the next node's IP address (43). The modified packet is then forwarded to the local ISP default router (44), and from there to the next node, along the selected path in the QVPN.

Fig. 5 illustrates what happens to an IP packet, which enters a termination node (51). The packet is forwarded to the IP level, where the current node IP address is unwrapped. Referring also to Fig. 11, the Qflow level compares the offset number to the current hop number, and since they match, this indicates that this node is the final/terminating node. Therefore, the packet is returned to its original format (52) and forwarded to the local ISP default router (53), and from there to the end user (55), through the ISP node (54).

Fig. 6A illustrates a functional overview of a Qnode. The key component is the Qflow application suite (executed on a server), which encompasses several applications that run above TCP/IP. The various applications handle the following tasks:

1. Qping – this is a process that creates and sends unidirectional Qping packets from one node to other nodes. The purpose of these Qping packets is to measure, calculate and determine the fastest packet transfer paths, between any two nodes.

2. Qreciever – this process determines whether a QFLOW packet is destined to the current node or to another node. If the packet is not destined to the current node, it is sent onwards to the next hop, in the VPN, as fast as possible. Otherwise it is forwarded for further processing within the current node.
3. Qcollector – a collection process that gathers the Qping packets at the final (i.e. terminating) node, and sends a Qresponse packet to the originating node.
4. Qdatain – the data handling process for originating end user data transfer.
5. Qdataout - the data handling process for terminating end user data transfer.
6. Qbest – a process that manages tables for the Qdata process and perhaps others.
7. Qlogger - a process that receives and logs into a file any message that was sent by any of the Qflow process.
8. Qstatistics - a process that accumulates statistics and OMs of the Qflow processes.

For a better understanding of the flow of different types of packets, from an originating end-user node to a terminating end user node, through the various processes, refer to figures 6B, 6C and 6D.

Referring to Fig. 6B, it illustrates the way that a Qping packet is handled, according to the preferred embodiment of the invention. The Qping thread sends a testing packet to the terminating node. The packet traverses the TCP/IP stack of the originating node and enters the network cable, on its way to the intermediate node. Upon arriving at the intermediate node, the Qflow increases the header's offset by one, wrap the packet with the IP address of the terminating node, and forwards the packet onwards. At the

terminating node, the packet traverses the TCP/IP stack, and sends it to the Qcollector process. The terminating node has to send back its response packet, and in order to do this, the terminating node turns into an originating node, and vice versa, as can be seen in Fig. 6C.

The intermediate node shown in Figures 6B and 6C is only for illustration purpose, and several nodes may be used as intermediate nodes, as was explained earlier.

Referring to Fig. 6C, it illustrates the flow process of the response packet. The Qcollector sends a packet back to the Qping originator. The packet first exits the (now) originating node and enters the intermediate node. Since the intermediate node recognizes, by identifying the Qflow header, that the packet still has to be forwarded to the (now) terminating node, its Qflow process increases the hopping offset number, and it also assigns to the packet the new terminating node IP address. At the terminating node, the packet is forwarded, by the Qflow process, to the Qbest process.

The originator node (in Fig. 6B) registers the time, at which the Qping packet is initiated/sent. Additionally, the time of its arrival at the Qcollector of the terminator node (in Fig. 6B) is registered in the packet, and extracted in the Qbest process, in the (now) terminating node (in Fig. 6C). Such sessions, of sending Qping packets to terminating node(s) and receiving the corresponding response(s) in the originating node(s), are carried out through different intermediate node(es), and their travel times are recorded in each originating node. This procedure allows the originating nodes to assign a QoS level to each potential route/path in the QVPN, from which several routes, which meet predetermined efficiency criteria, are selected.

Referring to Fig. 6D, it illustrates the way that a data packet is handled, according to the preferred embodiment of the invention. A packet from one end user arrives via the network to the originating node. The CAMOD driver sends the packet directly to the Qdatain process. Since the originating node receives the original packet, which contains the final end user IP address, the Qdatain process wraps the packet with the IP address of every intermediate Qnode in the selected QVPN path. The wrapped packet then enters the intermediate node, where the Qflow process updates the offset hop number and the next intermediate node IP address. The packet is then received in the terminating node identified by comparing the offset number with the current offset number. Since it is a terminating node, the packet is unwrapped from the last Qflow header, and is forwarded to the Qdataout, and from there to the recipient end user, through the local ISP .

The Qping testing guidelines and process:

The Qping concept enables the Qnodes to periodically measure all the possible paths, and to update their tables accordingly. Qpings are initiated periodically while the time-out between them can differ from one link/path to another, and/or during the time.

Fig. 7 is an example of the basic Qping concept. Several Qnodes (i.e. A, B, C, D, etc.) are connected to the Internet. For the sake of explanation, Qnodes A and E are the originator and terminator Qnodes, respectively. Several valid/selected paths may connect Qnodes A and E to each other. The number of intermediate nodes creating a preferred/valid path may differ in each path . A possible routing between points A and E can be carried out either directly (i.e. a direct path from point A to point E), in which case no intermediate Qnodes are used at all, or using one

intermediate Qnode (e.g. paths A-B-E, A-C-E). Other options are having two intermediate Qnodes (e.g. paths A-D-F-E, A-F-B-E) or even more. In other words, a valid/preferred path, between an originating node and a terminating node, may comprise any combination of nodes as intermediate nodes. Every said valid path is tested for its propagation time delay, in order to assess its quality. Although said QVPN uses the public Internet routers, these routers are transparent to this QVPN. The QVPN imposes an end-to-end routing policy on the Internet routers in the way that was explained before, namely by essentially using said intermediate Qnodes.

Fig. 8 shows an example of a private case of a Fig. 7, where there is only one intermediate node in each route/path in addition to the direct path (i.e. without any intermediate node) that connects the originating node A to the terminating node E. In this example (Fig. 8), the QVPN comprises only five possible/valid routes/paths, and a maximum of one hop. As was explained before, several intermediate nodes may be included in a QVPN path. These five above-mentioned paths are an example for paths that are predefined at the time of deploying the Qnodes. However, new paths can be defined later on, which may comprise combinations of already existing nodes and/or new installed Qnodes.

All of the valid paths, that the QVPN is comprised of, have to be measured for their 'quality'. This is carried out by sending "Qpings" (test packets) between every originator/source node and every terminating/destination node. The originating node A (see **Fig. 8**) sends Qping packets through the five paths to the terminating node E. Upon receiving these Qpings, the terminating node E measures the propagation time of the first received Qping and the time intervals between each subsequently received Qpings. All this information is sent back, as a response, to the source node A. Since node A registers the Qpings departure times and receives their arrival

times at E, by the Qresponse packet, the propagation time of each path is calculated and registered in a 'A-to-E' routing table. If a path, for instance ,via Qnode C is extremely congested, the appropriate Qping packet may not reach the terminating node E in which case this packet is recorded as a lost packet. Consequently, such a path shall not be used for data transmission. Node E (the destination) ends the Qping evaluation process whenever at least one of the following three events occurs:

- all Qpings are received (the total number of Qpings sent by node A is inserted in each Qping packet);
- after a time-out from receipt of first Qping;
- a new Qping session is received by E.

At the end of the Qping reception process, node E (the terminating node) sends a response packet to the originator node A; i.e. a Qresponse packet, which carries the measurements results. Upon receipt of the Qresponse packet at the originator node A, the originating node initiates an evaluation process of the various possible paths to destination, based on the information contained in the Qresponse packet, but also on additional data like:

- historical data (e.g. lost packets, delay jitter).
- preset conditions inserted via the QVPN Network Management like: costs factors associated to various paths, limitations in using several intermediate Qnodes, etc.
- type of data (voice, video, file transfer etc.)

At the end of the evaluation process, node A associates a quality factor to each path selected for the data transmission. Different quality factors are assigned to different types of traffic, applications (e.g. voice, data, video) and customers. The combination of said measured paths' quality with the application's parameters determines the paths' weight. In addition to the

weight factor, other predefined parameters/factors are involved in choosing the optimized paths. The higher the weight of a path(s), the better the path(s) suits a specific application or user. It should be noted that each path might have different weights at the same time, since a path may be considered as an adequate path for data packets, thus having a relatively high weight, while the same path may be considered as a poor path for voice packets, thus having a low weight.

There are several options, by which data packets may be transferred from node A (i.e. the originator node) to node E (i.e. the terminating node). One option is via one preferred/optimized path. Another option is to transfer data packets via two or more paths (i.e. a multi-path data transfer). Node A may decide to apply the multi-path option in order to implement load balancing. The multi-path option may be utilized in various ways, one of which is allocating several specific paths to one specific application at a time, providing that each path has adequate weight. Another way to utilize the multi-path option is, that several applications share some, or all, of the paths; namely paths are used intermittently by several applications. Referring again to Fig. 8, the originating node A may decide, for example, to start sending voice packets to point E through nodes B and C, and data packets through nodes D and F. However, at a later stage node A may decide, for example to divert voice packets from node C to node F, and data packets from node F to node C. These kind of changes are made dynamically by the originator node A, and are based essentially on updated weights which are assigned to each path.

Fig. 9 illustrates the structure of the Qping packet. The fields in this type of packet are:

1. 'OP CODE' – Is an operation code that indicates that this is a Qping packet.
2. 'TOTAL LENGTH' – Is the length of the Qping packet, excluding the 'OP CODE' and length fields.
3. 'HEADER LENGTH' - Is the length of the header in 'words'.
4. 'QOS' – It is a flag that indicates the Quality of Service level that is assigned to this packet.
5. 'NUMBER OF HOPS' – the total number of nodes, through which the packet is to be forwarded, from the originating node to the terminating node.
6. 'HOPS OFFSET' – a counter that counts the current number of Qnodes, that the packet has already traveled through. In other words, each time the packet reaches the next Qnode, this counter is incremented by one.
7. 'HOP 1 ADDRESS' – the address of the first Qnode, the packet is forwarded to.
8. all additional intermediate Qnodes addresses in the order of locations along the path.
9. 'HOP n ADDRESS' – the address of the termination /node.
10. 'ORIGINATING ADDRESS' – it's the IP address of the node that created this Qping packet. It is required in order for the termination Qnode to know where to send the Qresponse packet, namely to send a response back to the initiating (i.e. originating) Qnode.
11. 'TOTAL PATH' – the total number of paths/, in the QVPN, through which the originating Qnode sends Qping to the terminating Qnode. This number is used by the terminating Qnode to determine when a Qping session is completed, so that it would not have to wait for other Qping to arrive. After having received all the expected Qpings (from a specific originating Qnode), the terminating Qnode starts to send response (i.e. Qresponse) back to the originating Qnode.

12. 'PATH NUMBER' – this is the path number of this Qping packet. This number is required in order for the originating Qnode to match the propagation (delay) time to the corresponding path, so that it can assign the right quality to the right path.
13. 'SESSION NUMBER' – this is the session number of this Qping.
14. 'Start Time [seconds]' – the time (in seconds) that this packet was sent from the originating node.
15. 'START TIME [miliseconds]' – the time (in miliseconds) that this packet was sent from the originating node.
16. 'OPTIONAL DATA PADDING' – optional number of bytes, to simulate packets of various sizes.

Qping packets with the same format are sent along each one of predefined QVPN's paths from every originating Qnode to every terminating node. The decision, to be taken by an originating node, when and with whom to start a Qping session, is determined on the basis of efficiency. Preferably, active paths will be sampled (i.e. checked) frequently, while non-active paths will not be sampled at all, in order not to interfere (i.e. not to load) with the normal operation of the Internet network. However, other modes of Qping/sampling sessions, that comply with the limits of the efficiency and network congestion criteria are possible too.

After all of the Qpings, that were sent from an originating node to a terminating node, arrive at the terminating node, the terminating node starts sending back the corresponding response.

Fig. 10 illustrates the structure of the Qresponse packet. The fields in this type of packet are similar to the fields of the Qping packet, excluding the following fields:

1. 'OP CODE' – Is an operation code that indicates that this is a Qresponse packet.
2. ORIGINATOR ADDRESS' – this is the IP address of the current node, that is sending this Qresponse packet. This address allows the originating Qnode to identify the source of received Qresponse packet so that it can match a specific route to a specific terminating node.
3. 'PACKET TIME STRUCTURE 1' – this raw, in the table, contains the path number and the propagation time of the fastest path. The better the path, the smaller this number.
4. 'PACKET TIME STRUCTURE n' – contains the path number and the delta time between the fastest path and path number 'n', where 'n' is the number of paths

After the originating node measures the time delay to the terminating node, over the predefined paths, it selects the most appropriate paths for each data packet.

Fig. 11 illustrates the structure of the Qdata packet. The fields in this type of packet are similar to the fields of the Qping packet, excluding the following fields:

1. 'ORIGINAL END USER PACKET' – this is the original end user data packet. All of the previous fields, from the 'Op Code' to the 'Hop n Address', are added to the original data packet by the originating node (i.e. Qflow application), in order to forward the data packet across and over the QVPN. The routers of the Internet network do not have access to the real final destination (i.e. end user's IP), but only to the next intermediate Qnode, as is reflected in the modified data packet (see Fig. 3 for wrapping the original Qdata packet, 33, with a new header 32).

The above examples and description have of course been provided only for the purpose of illustration, and are not intended to limit the invention in any way. As will be appreciated by the skilled person, the invention can be carried out in a great variety of ways, employing more than one technique from those described above, all without exceeding the scope of the invention.

CLAIMS

1. A method for improving the quality of transportation of selected data packets over a data network, comprising:
   a) determining selected nodes being access points to said data network, each of which may be a source from which said selected data packets can be sent, or a destination to which said selected data packets can be intended;
   b) selecting a plurality of intermediate nodes between said source and said destination, for generating a plurality of alternative paths, consisting of segments, for routing said selected data packets;
   c) periodically sampling the packet transportation parameters in the segments of each preselected path, each time by sending a plurality of test packets from said source to said destination, along said preselected different paths defined by different intermediate nodes and their corresponding interconnecting segments;
   d) defining one or more optimal paths for delivering said selected data packets from said source to said destination according to said transportation parameters and optionally, also according to predefined parameters characterizing said segments by selecting a combination of segments, connected to nodes, and having the optimal sampled transportation parameters and/or predefined parameters, that connects said source to said destination;
   e) for each selected data packet, generating a modified header containing a sequence of consecutive addresses that correspond to consecutive nodes along an optimal path and attaching said modified header to said selected data packet;
   f) forwarding each selected data packet from said source to said destination along said optimal path, while at each node along said optimal path, starting from the source:

f.1) processing said modified header;

f.2) extracting the address that corresponds to the next consecutive node;

f.3) forwarding said selected data packet from said node to its consecutive node using the extracted address;

f.4) repeating steps f.1) to f.3) for all intermediate nodes until said destination node; and

g) at the destination node, removing said modified header from said selected data packet and whenever desired, allowing using its original header.

2. A method according to claim 1, wherein the data network is the Internet.

3. A method according to claim 1, wherein one or more nodes are used as intermediate nodes.

4. A method according to claim 1, wherein the test packet does not contain a payload.

5. A method according to claim 1, wherein the transportation parameters are selected from the following group of parameters:

- the delay time of data packets from source to destination;
- the variation of said delay time; and
- loss of packets.

6. A method according to claim 1, wherein data is concurrently delivered from a source to a destination over several paths.

7. A method according to claim 6, further comprising using weighted distribution of data between paths.

8. A method according to claim 7, wherein the weighted distribution is determined according to the desired level of QoS between the source and the destination.

9. A method according to claim 4, wherein the definition of the optimal path is carried out by measuring and storing the time and/or the order of arrival of test packets through different paths from the source to the destination.

10. A method according to claim 1, further comprising:
    a) dynamically varying the definition of each optimal path from the source to the destination, according to the sampling results; and
    b) whenever a new optimal path is defined, continuing sending data packets from said source to said destination over said new optimal path.

11. A method according to claim 1, wherein the optimal path consists of direct connection between the source and the destination.

12. A method according to claim 1, wherein the predefined parameters are selected from the following group of parameters:

- cost;
- availability;
- agreements with ISPs;
- data type;
- agreement with customers.

13. A method according to claim 1, wherein a grade is assigned to each optimal path according to the sampled transportation parameters and/or predefined parameters that correspond to a required QoS and/or to the type of data packets to be sent from the source to the destination.

14. A method according to claim 13, further comprising dynamically varying the grade of at least one optimal path according to the sampled transportation parameters and/or to the type of data packets to be sent from the source to the destination.

15. A method according to claim 13, wherein voice packets are sent from the source to the destination through one or more optimal paths being optimal for voice.

16. A method according to claim 13, wherein data packets are sent from the source to the destination through one or more optimal paths being optimal for data.

17. A method according to claim 13, further comprising splitting the transportation of data packets from the source to the destination through between two or more optimal paths, such that more transportation is directed to, and distributed between, optimal paths having higher grades than the remaining optimal paths, and less transportation is directed to, and distributed between, said remaining optimal paths.

18. A data network having improved quality of transportation of selected data packets, comprising:

a) a plurality of nodes being access points to said data network, each of which may be a source from which said selected data packets can be

sent, or a destination to which said selected data packets can be intended;

b) a plurality of intermediate nodes between said source and said destination, for generating a plurality of alternative paths, consisting of segments, for routing said selected data packets;

c) at one or more nodes and/or intermediate nodes, circuitry for sending a plurality of test packets from said source to said destination, along said preselected different paths defined by different intermediate nodes and their corresponding interconnecting segments;

d) processing means for defining one or more optimal paths for delivering said selected data packets from said source to said destination according to said transportation parameters and optionally, also according to predefined parameters characterizing said segments, and for selecting a combination of segments, connected to nodes, and having the optimal sampled transportation parameters and/or predefined parameters, that connects said source to said destination;

e) at each source, processing means for generating a modified header, for each selected data packet, that contains a sequence of consecutive addresses that correspond to consecutive nodes along an optimal path and attaching said modified header to said selected data packet;

f) at each node along said optimal path, starting from the source:

f.1) processing means for processing said modified header and for extracting the address that corresponds to the next consecutive node;

f.2) circuitry for forwarding said selected data packet from said node to its consecutive node along said optimal path using the extracted address; and

g) at the destination node, processing means for removing said modified header from said selected data packet and for obtaining the original header of said selected data packet.

19. A data network according to claim 18, being the Internet.

20. A data network according to claim 18, in which one or more nodes are used as intermediate nodes.

21. A data network according to claim 18, in which the test packet does not contain a payload.

22. A data network according to claim 18, in which the transportation parameters are selected from the following group of parameters:

- the delay time of data packets from source to destination;
- the variation of said delay time; and
- loss of packets.

23. A data network according to claim 18, in which data is concurrently delivered from a source to a destination over several paths.

24. A data network according to claim 23, comprising weighted distribution of data between paths.

25. A data network according to claim 24, in which the weighted distribution is determined according to the desired level of QoS between the source and the destination.

26. A data network according to claim 4, in which the definition of the optimal path is carried out by measuring and storing the time and/or

the order of arrival of test packets through different paths from the source to the destination.

27. A data network according to claim 18, further comprising:
processing means for dynamically varying the definition of each optimal path from the source to the destination, according to the sampling results and for sending data packets from said source to said destination over said new optimal path.

28. A data network according to claim 18, in which the optimal path consists of direct connection between the source and the destination.

29. A data network according to claim 18, in which the predefined parameters are selected from the following group of parameters:

- cost;
- availability;
- agreements with ISPs;
- data type;
- agreement with customers.

30. A data network according to claim 18, in which a grade is assigned to each optimal path according to the sampled transportation parameters and/or predefined parameters that correspond to a required QoS and/or to the type of data packets to be sent from the source to the destination.

31. A data network according to claim 30, that comprises processing means for dynamically varying the grade of at least one optimal path according to the sampled transportation parameters and/or to the type of data packets to be sent from the source to the destination.

32. A data network according to claim 30, in which voice packets are sent from the source to the destination through one or more optimal paths being optimal for voice.

33. A data network according to claim 30, wherein data packets are sent from the source to the destination through one or more optimal paths being optimal for data.

34. A data network according to claim 30, comprising processing means for splitting the transportation of data packets from the source to the destination through between two or more optimal paths, such that more transportation is directed to, and distributed between, optimal paths having higher grades than the remaining optimal paths, and less transportation is directed to, and distributed between, said remaining optimal paths.

35. A method for improving the quality of transportation of selected data packets over a data network, substantially as described and illustrated.

36. A data network having improved quality of transportation of selected data packets, substantially as described and illustrated.

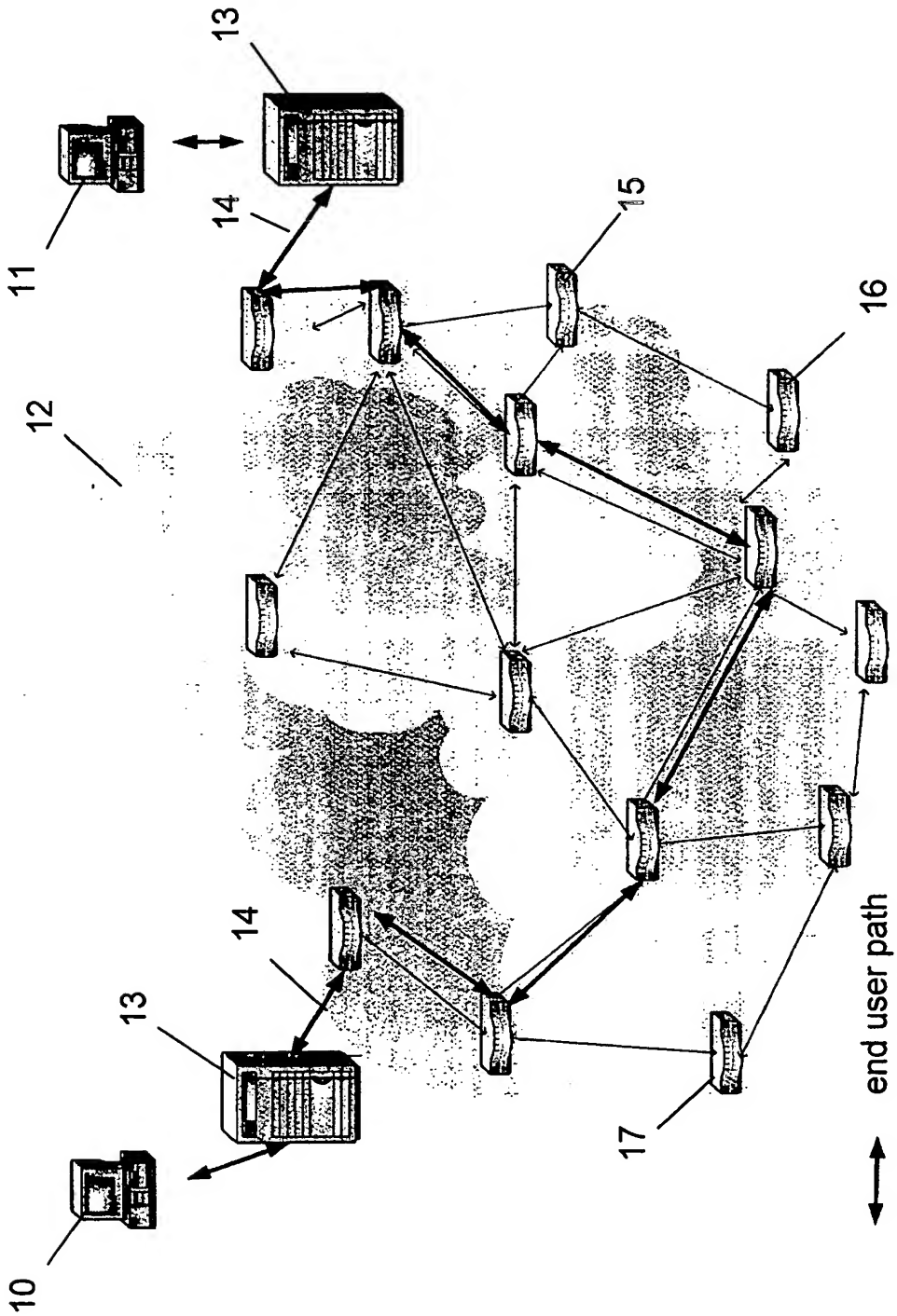


Fig. 1 (prior art)

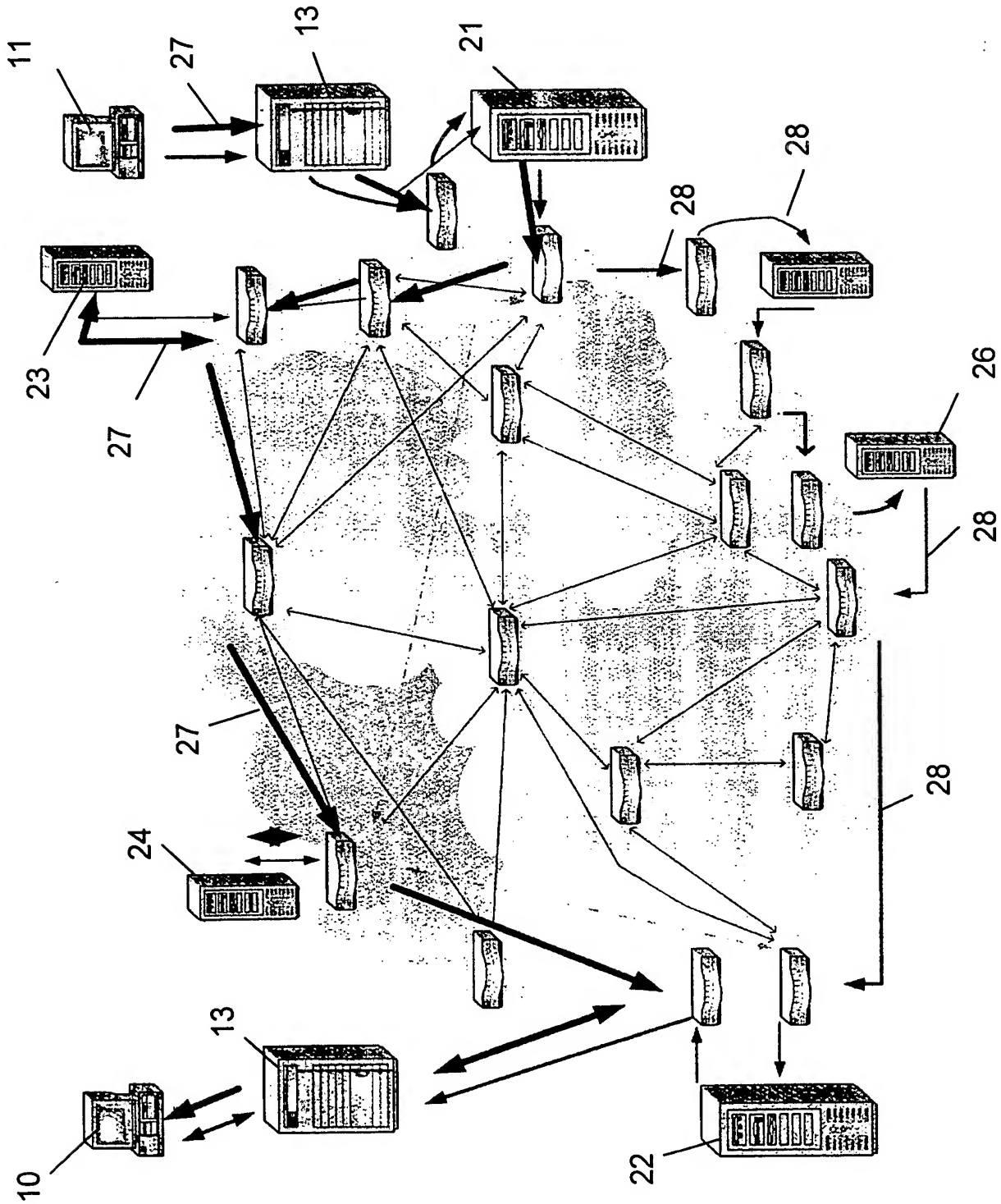


Fig. 2

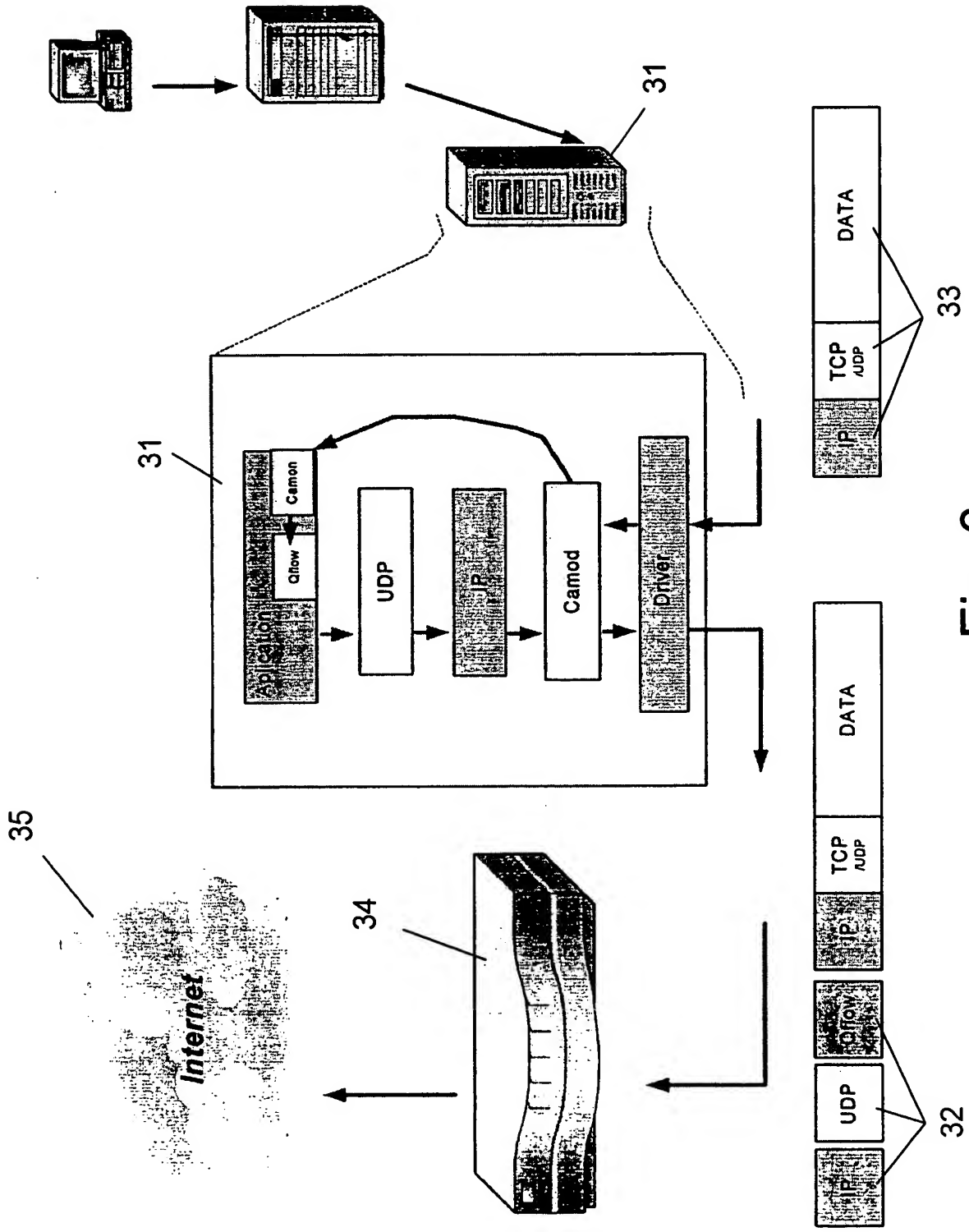

31
31
Application
Camon
Qflow
UDP
IP
Camod
Driver
DATA
TCP
UDP
IP
33
DATA
TCP
UDP
IP
Qflow
UDP
IP
32
34
35
Internet

Fig. 3

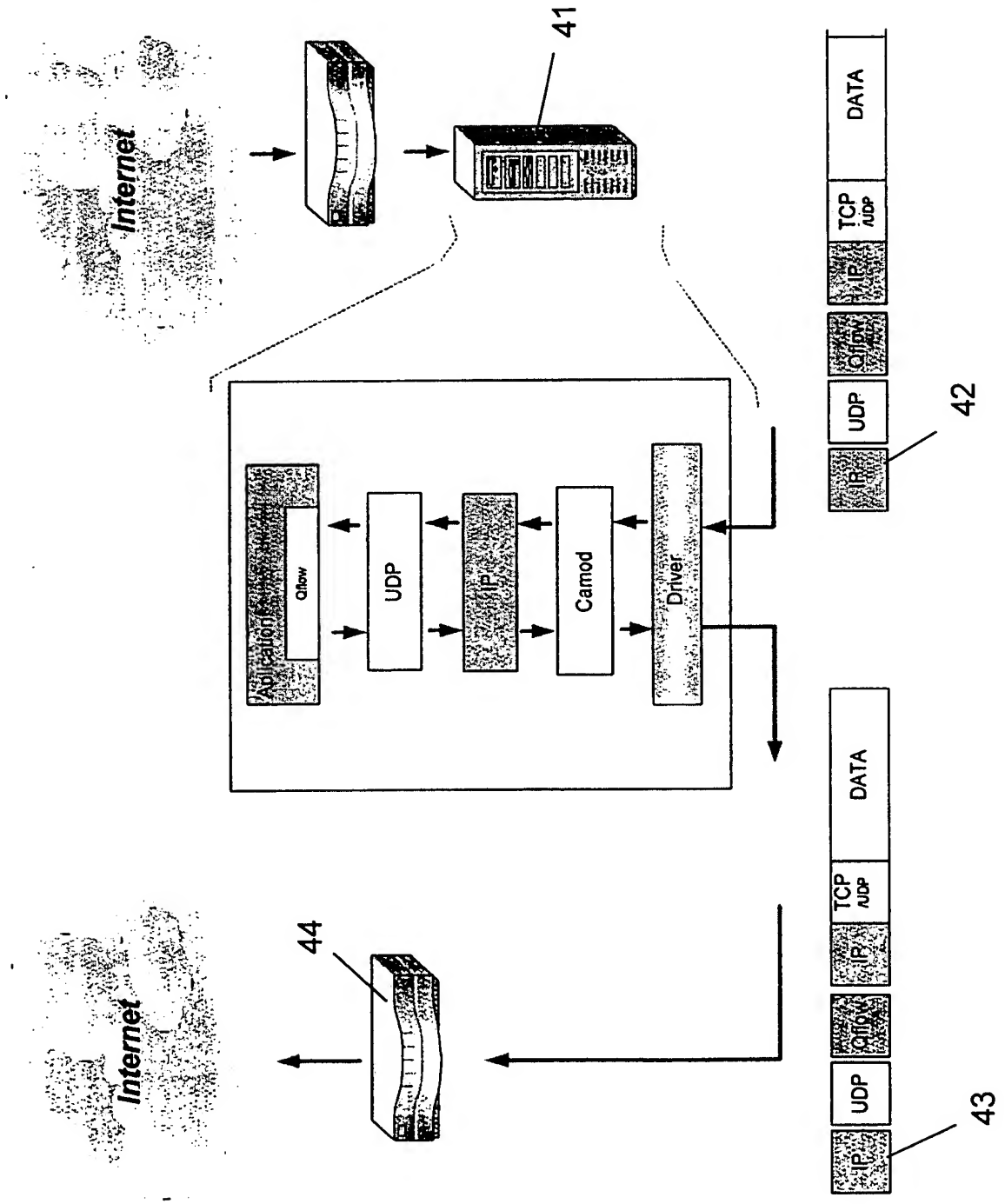

Internet
41
Application
Qflow
UDP
IP
Camod
Driver
IP
UDP
Qflow
TCP
DATA
42
Internet
44
IP
UDP
Qflow
IP
TCP
DATA
43

Fig. 4

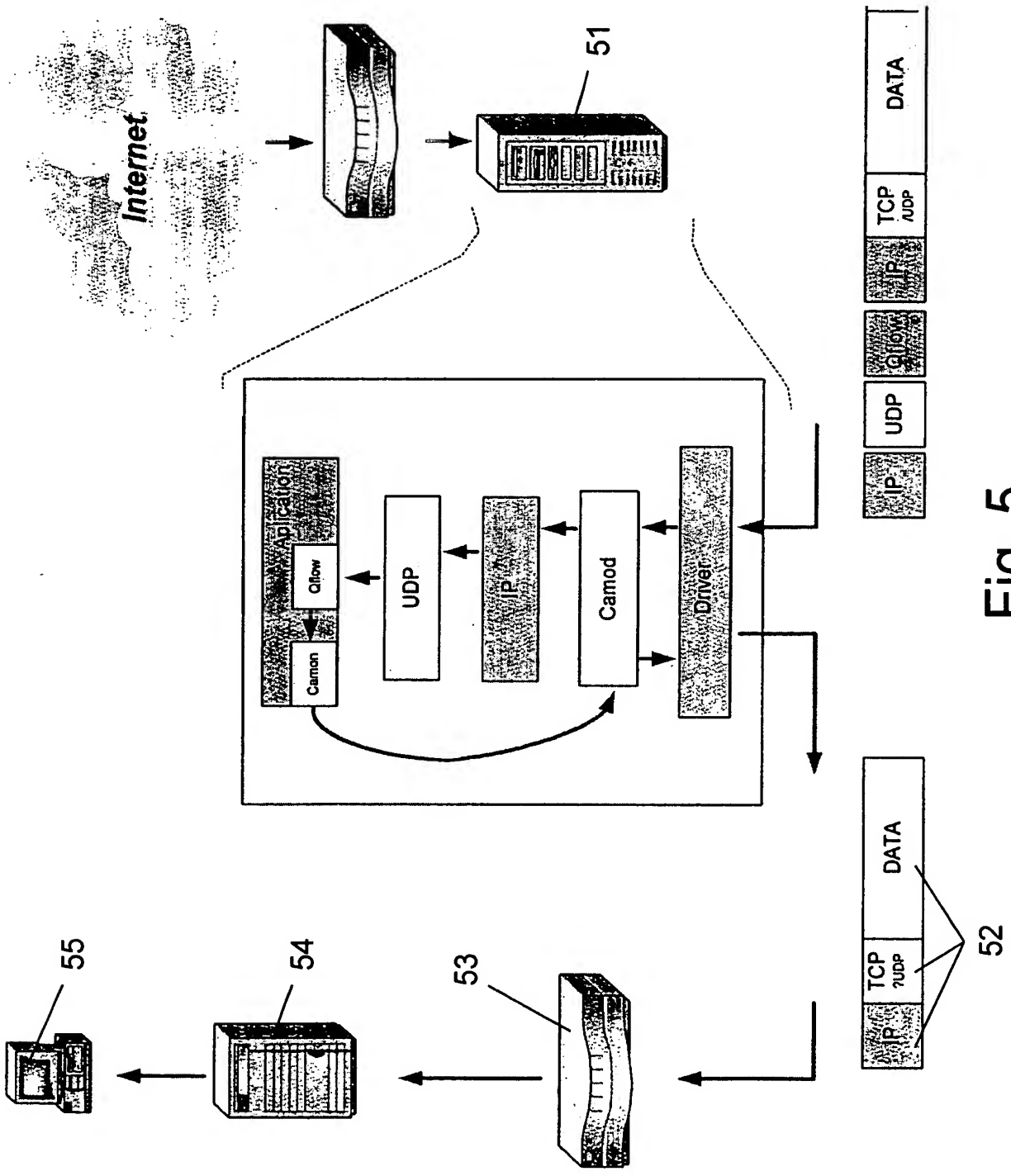

Internet
51
Application
Camon
Qflow
UDP
IP
Camod
Driver
IP
UDP
Qflow
TCP /UDP
DATA
IP
TCP /UDP
DATA
52
53
54
55

Fig. 5

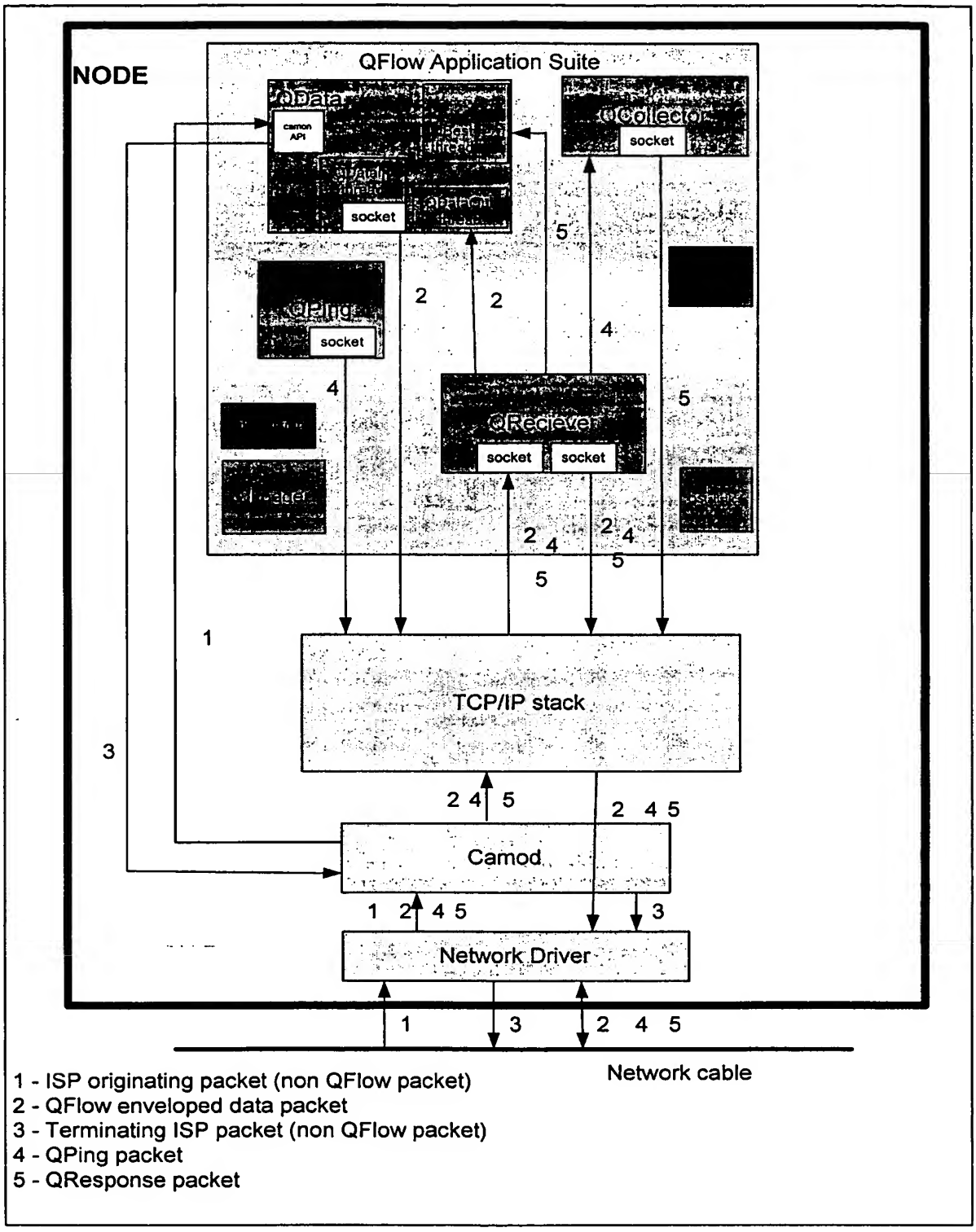

NODE
QFlow Application Suite
QData
socket
QCollector
socket
QPing
socket
QRecieve
socket
socket
TCP/IP stack
Camod
Network Driver
Network cable
1 - ISP originating packet (non QFlow packet)
2 - QFlow enveloped data packet
3 - Terminating ISP packet (non QFlow packet)
4 - QPing packet
5 - QResponse packet

Fig. 6A

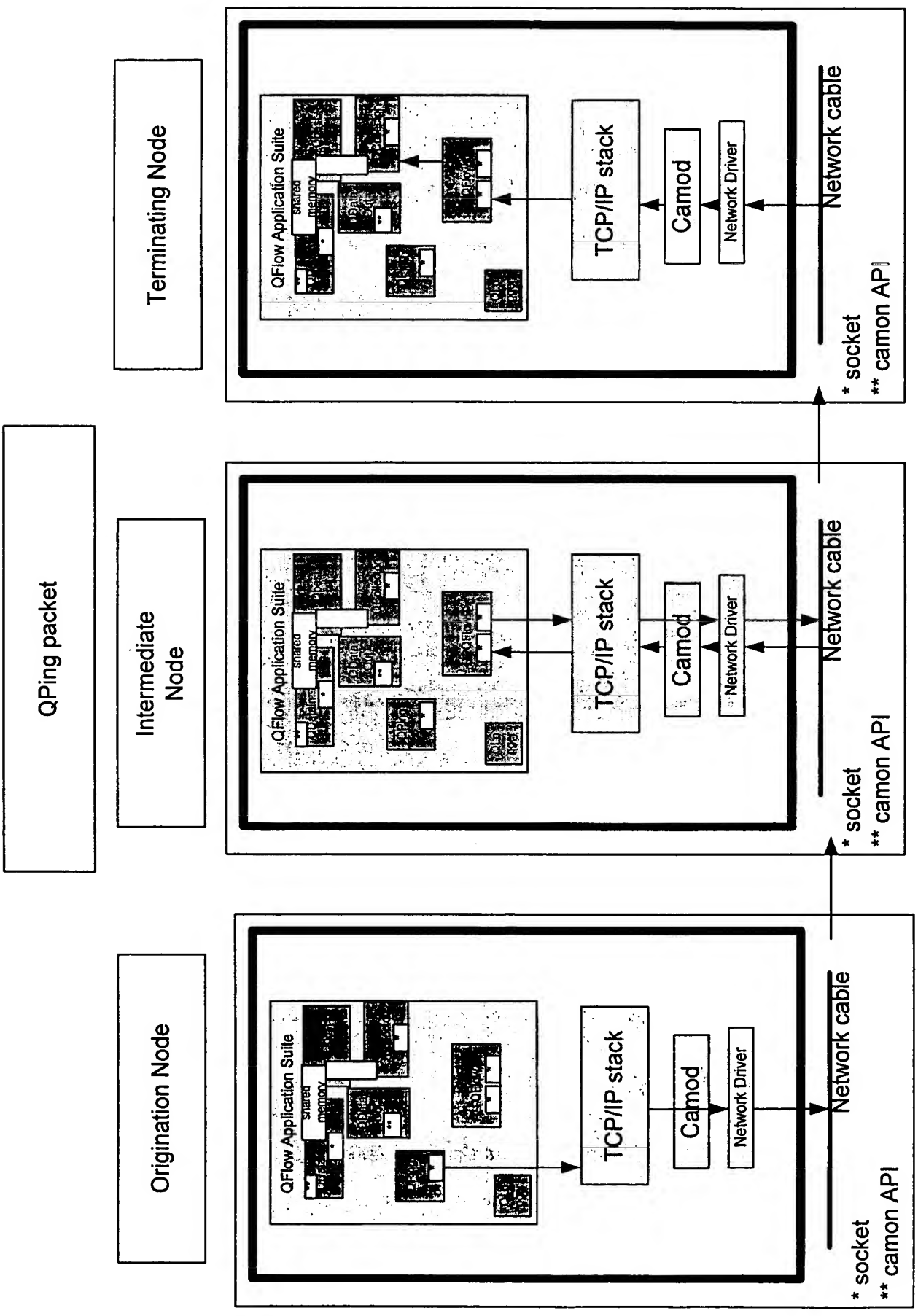

QPing packet
Origination Node
Intermediate Node
Terminating Node
QFlow Application Suite
shared memory
TCP/IP stack
Camod
Network Driver
Network cable
* socket
** camon API

Fig. 6B

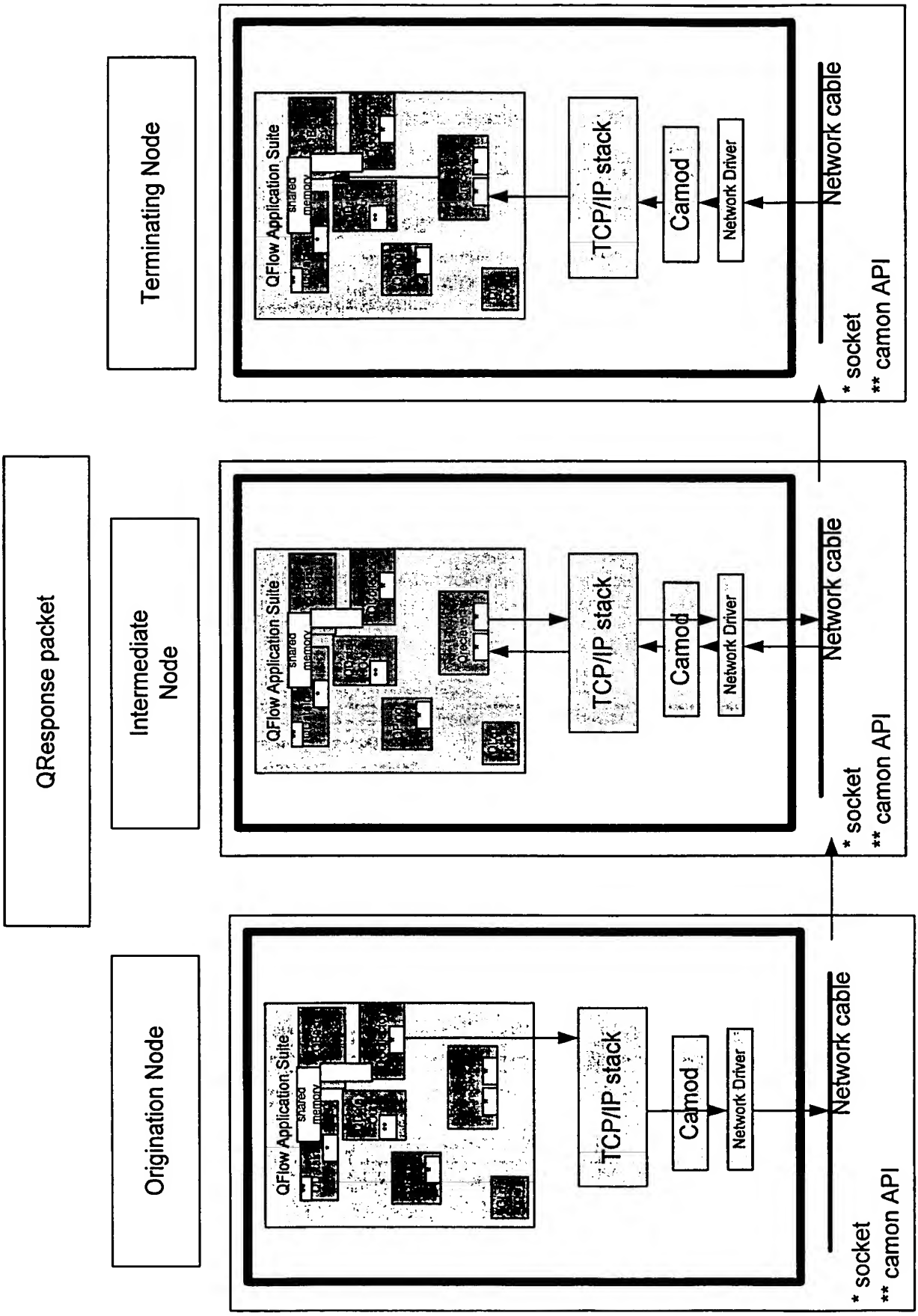

QResponse packet
Origination Node
Intermediate Node
Terminating Node
QFlow Application Suite
shared memory
TCP/IP stack
Camod
Network Driver
Network cable
* socket
** camon API

Fig. 6C

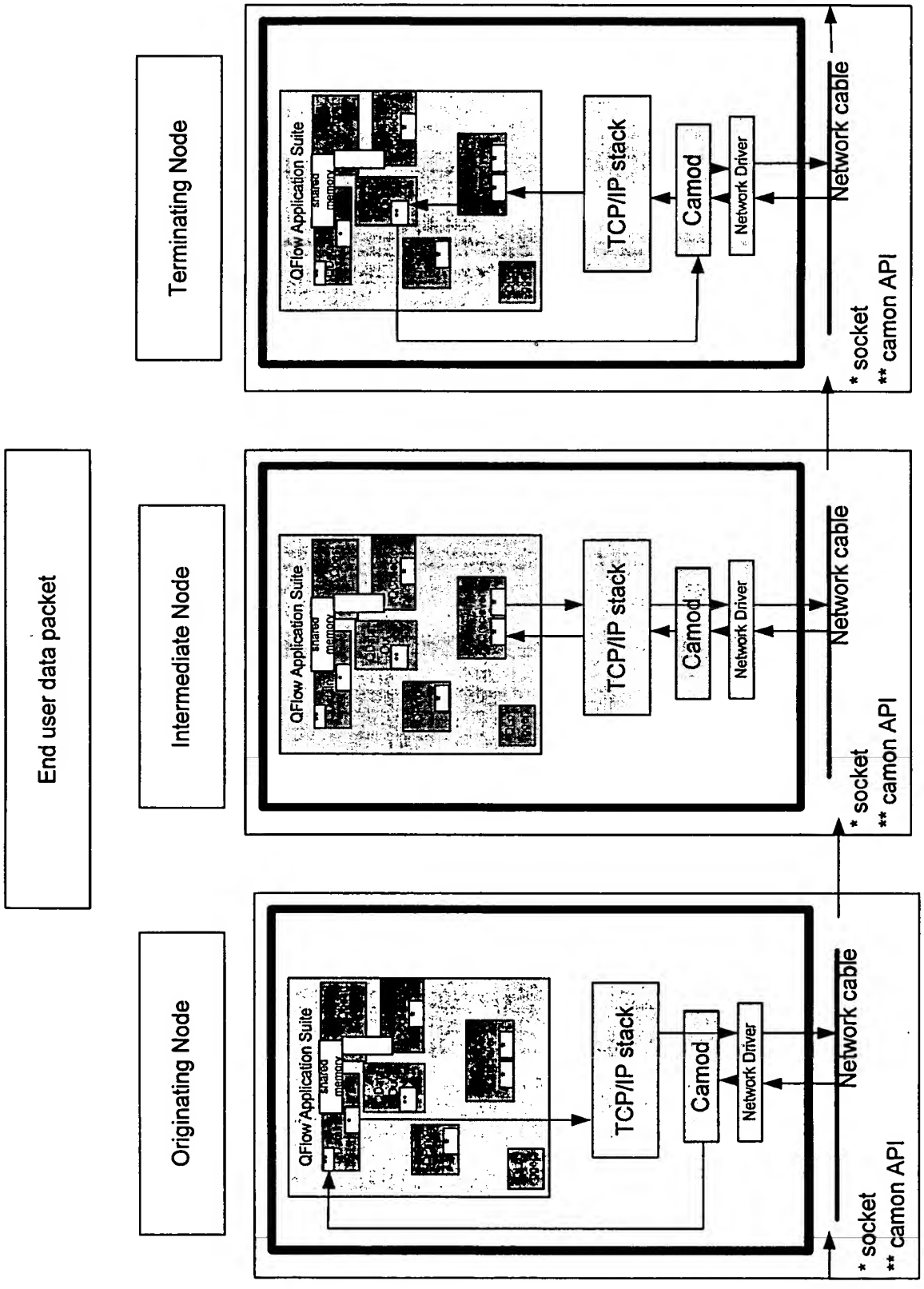

End user data packet
Originating Node
Intermediate Node
Terminating Node
QFlow Application Suite
shared memory
TCP/IP stack
Camod
Network Driver
Network cable
* socket
** camon API

Fig. 6D

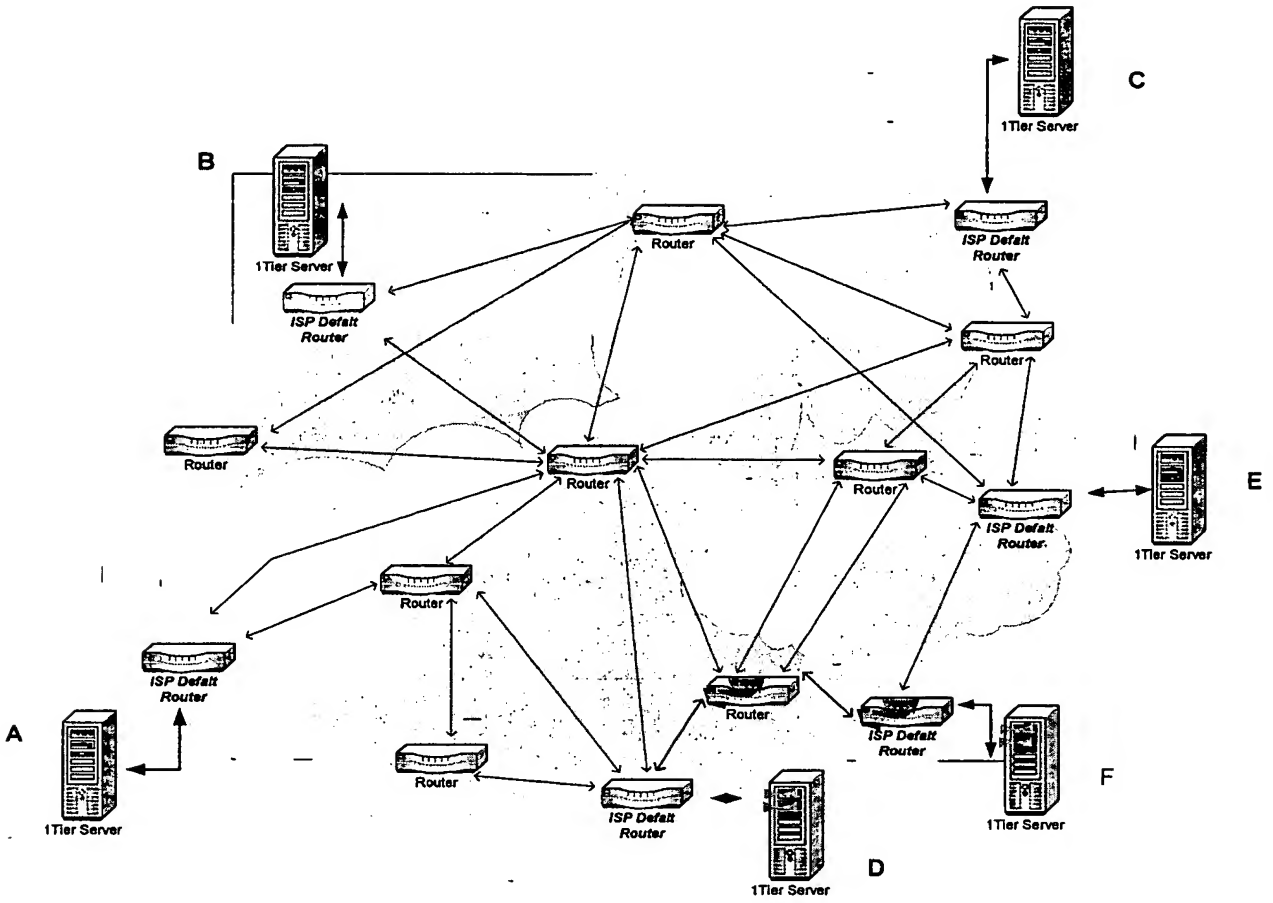


Fig. 7

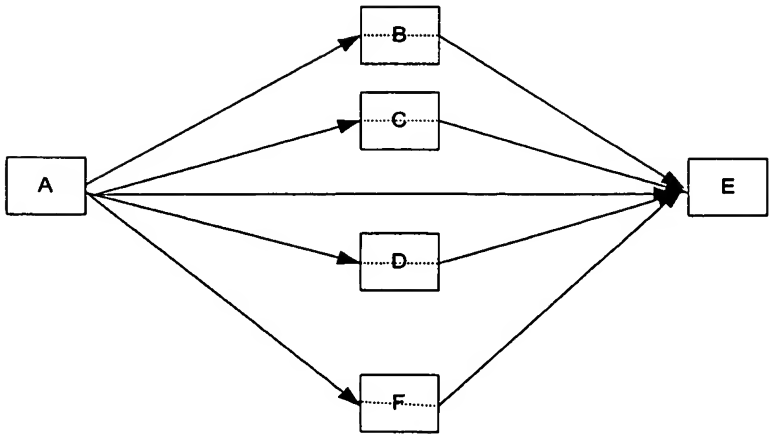


Fig. 8

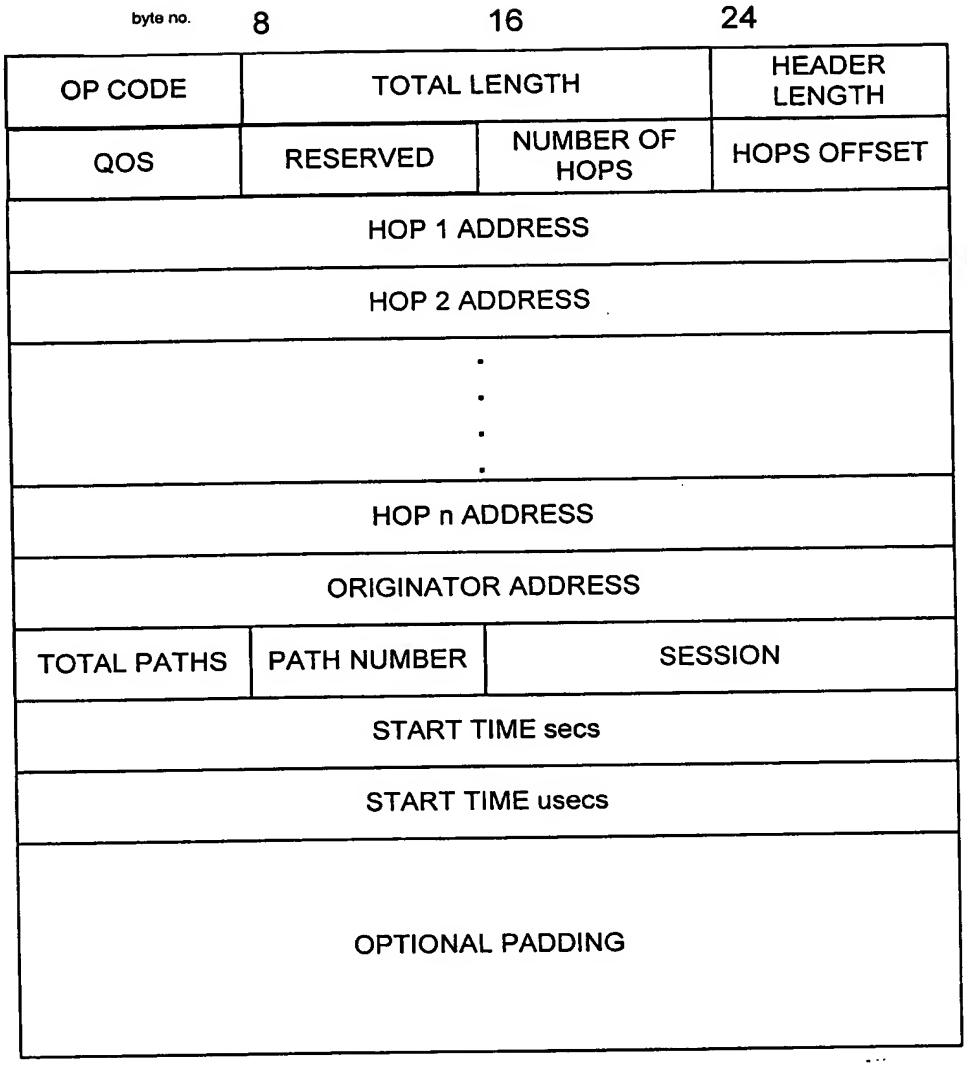

byte no.
8
16
24
OP CODE
TOTAL LENGTH
HEADER LENGTH
QOS
RESERVED
NUMBER OF HOPS
HOPS OFFSET
HOP 1 ADDRESS
HOP 2 ADDRESS
HOP n ADDRESS
ORIGINATOR ADDRESS
TOTAL PATHS
PATH NUMBER
SESSION
START TIME secs
START TIME usecs
OPTIONAL PADDING

Fig. 9

byte no. 8	16	24
OP CODE	TOTAL LENGTH	HEADER LENGTH
QOS	RESERVED / NUMBER OF HOPS	HOPS OFFSET
HOP 1 ADDRESS		
HOP 2 ADDRESS		
. . . .		
HOP n ADDRESS		
ORIGINATOR ADDRESS		
TOTAL PATHS	RESERVED	SESSION
PATH 1	DELTA TIME ms	EMPTY
PATH 2	DELTA TIME ms	EMPTY
. . . .		
PATH n	DELTA TIME ms	EMPTY

Fig. 10

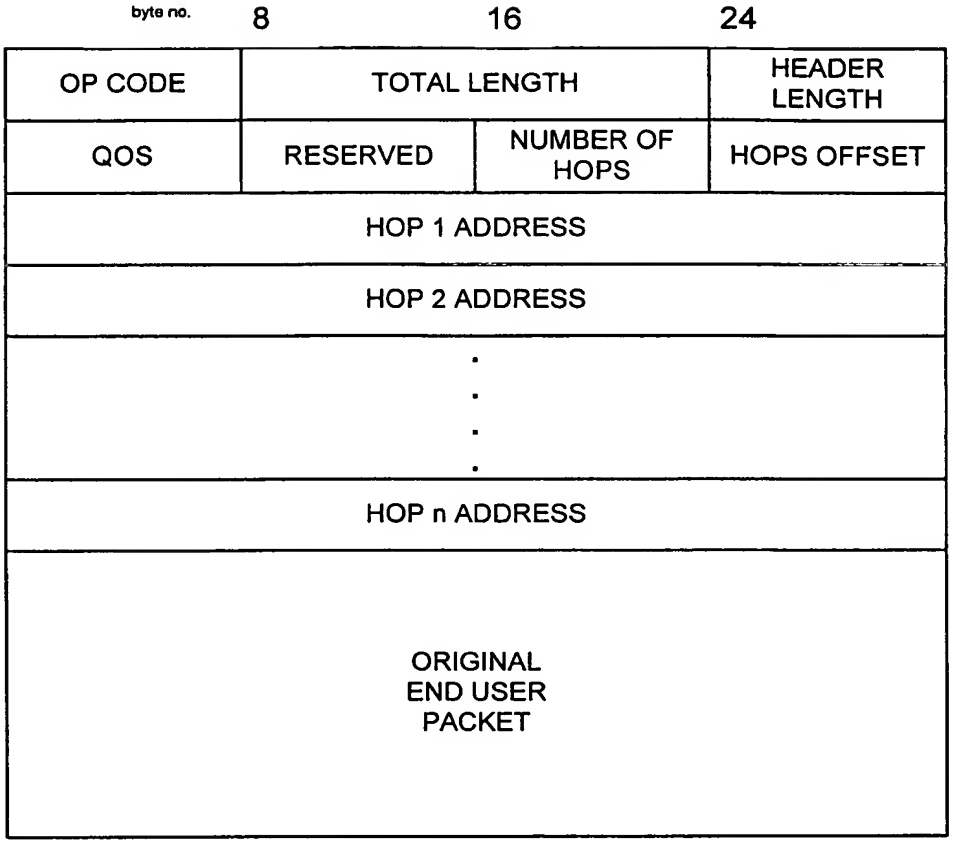

byte no.
8
16
24
OP CODE
TOTAL LENGTH
HEADER LENGTH
QOS
RESERVED
NUMBER OF HOPS
HOPS OFFSET
HOP 1 ADDRESS
HOP 2 ADDRESS
HOP n ADDRESS
ORIGINAL END USER PACKET

Fig. 11